Slepian-Bangs-type formulas and the related Misspecified Cramér-Rao Bounds for Complex Elliptically Symmetric distributions

Abdelmalek Mennad, Stefano Fortunati, Mohammed Nabil El Korso, Arezki Younsi, Abdelhak M. Zoubir, Alexandre Renaux

Slepian-Bangs-type formulas and the related Misspecified Cramér-Rao Bounds for Complex Elliptically Symmetric distributions

Abdelmalek Mennad[a], Stefano Fortunati[b,d,*], Mohammed Nabil El Korso[c], Arezki Younsi[a], Abdelhak M. Zoubir[d], Alexandre Renaux[e]

[a] *Laboratoire Micro-Onde et Radar, Ecole Militaire Polytechnique, P.O. Box 17, Bordj El Bahri 16111, Algeria*
[b] *Dipartimento di Ingegneria dell'Informazione, Univesity of Pisa, Pisa 56122, Italy*
[c] *Paris Ouest University, LEME EA4416, 50 rue de Sèvres, Ville d'Avray 92410, France*
[d] *Signal Processing Group, Technische Universität Darmstadt Merckstr. 25, Darmstadt 64283, Germany*
[e] *Université Paris-Sud, Orsay, Dpt Physics Laboratory of Signals and Systems (L2S) Modelisation and Estimation Group, France*

1. Introduction

The asymptotic performance analysis of an estimation algorithm mostly relies on two simplified assumptions: *i*) the data are assumed to be Gaussian distributed and *ii*) the data model used to derive the estimation algorithm is supposed to be *correctly specified*, that is the probability density function (pdf) assumed to derive an estimator of the parameters of interest and the true pdf that statistically characterizes the data are exactly the same.

Although these assumptions guarantee the possibility to perform "elegant" performance assessment, e.g. by evaluating the Cramér-Rao Bound (CRB) for the estimation problem at hand and/or by obtaining a closed form expression for the Mean Square Error (MSE) of a given estimator, the everyday engineering practice clearly calls the hypotheses *i*) and *ii*) into question. Regarding the Gaussian model assumption, large-scale measurement campaigns and the subsequent statistical analysis of the data gathered from a plethora of engineering applications, e.g. outdoor/indoor mobile communications, radar/sonar systems or magnetic resonance imaging (MRI), have highlighted the impulsive, heavy-tailed behaviour of the observations [1]. These experimental evidences have motivated the need to go beyond the Gaussian model and develop new statistical models able to better characterize the data. One of the more flexible and general non-Gaussian model is represented by the set of the Complex Elliptically Symmetric (CES) distributions [2], also called Multivariate Elliptically Contoured distributions [3]. CES distributions encompasses the complex Gaussian, the Generalized Gaussian and all the Compound Gaussian (CG) distributions, such as the complex *t*-distribution and the *K*-distribution, as special cases. The pdf of a CES distributed N-dimensional random vector $\mathbf{x}_l \in \mathbb{C}^N$ is completely characterized by the mean value $\boldsymbol{\gamma}$, the scatter (or shape) matrix $\boldsymbol{\Pi}$ and by a real valued function $w(t) : \mathbb{R}^+ \to \mathbb{R}$, called the *density generator*, i.e. $\mathbf{x}_l \sim CES_N(\boldsymbol{\gamma}, \boldsymbol{\Pi}, w)$ [2,3]. The CES distributions have been used in a variety of applications, in particular in the radar and array signal processing fields.

Other experimental evidences reveal recurring violations of the matched model assumption, that is the claim of a perfect match between the assumed and the true data model. The mathematical bases of a formal theory of the parameter estimation under model misspecification has been firstly developed by statisticians as Huber [4], White [5] and Vuong [6] and recently rediscovered by the Signal Processing (SP) community [7–9] and applied to a va-

* Corresponding author at: Dipartimento di Ingegneria dell'Informazione, Univesity of Pisa, 56122, Pisa, Italy.
E-mail address: stefano.fortunati@iet.unipi.it (S. Fortunati).

riety of well-known engineering problems: to Direction-of-Arrival (DoA) estimation in array and MIMO processing [7,10], to covariance/scatter matrix estimation in CES distributed data [8,11,12], to radar-communication systems coexistence [13] and to waveform parameter estimation in the presence of uncertainty in the propagation model [14], just to name a few.

This brief discussion clearly highlights the need to overtake both the Gaussian and the matched model assumptions while assessing the (asymptotic) performance of an estimator. As extensively discussed in the SP literature, one of the main tool for the performance assessment is the CRB that provides a lower bound to the MSE achievable by any unbiased estimator for a given estimation problem (see e.g. [15]). Under the matched model assumption, the CRB can be evaluated as the inverse of the Fisher Information Matrix (FIM), then having a convenient and easy way to evaluate the FIM would be of great practical utility. To this end, in array processing applications, the celebrated Slepian-Bangs (SB) formula has been introduced. Developed in the seminal works of Slepian [16] and Bangs [17], the SB formula provides a useful and compact expression of the FIM for vector parameter estimation under both Gaussian and matched model assumptions [15, Chapter 3, Appendix 3C]. Specifically, let $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$ be a d-dimensional, deterministic parameter vector and let $\mathbf{x} = \{\mathbf{x}_l\}_{l=1}^{L}$ with $\mathbf{x}_l \in \mathbb{C}^N$, be a set of L *independent* (possibly) complex random vectors, usually called *snapshots*, representing the available observations. If we assume that each snapshot follows a (complex) Gaussian parametric model, such that $\mathbf{x}_l \sim \mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}))$, then the FIM for the estimation of $\boldsymbol{\theta} \in \Theta$ can be expressed by means of the SB formula.

The first generalization of the SB formula to a non-Gaussian, but still perfectly matched, data model has been proposed by Besson and Abramovich in [18]. Specifically, Besson and Abramovich derived a compact expression for the FIM for the estimation of $\boldsymbol{\theta} \in \Theta$ when each snapshot $\mathbf{x}_l$ is characterized by a parametric CES distribution, i.e. $\mathbf{x}_l \sim CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}),\ \boldsymbol{\Pi}(\boldsymbol{\theta}),\ w)$. Note that the functional form of the parametrized mean value $\boldsymbol{\gamma}_l(\boldsymbol{\theta})$ is allowed to vary from snapshot to snapshot, while the covariance matrix is assumed to be constant. Clearly, since the Gaussian model belongs to the CES class, this generalized SB formula collapses to the classical one if the data are Gaussian distributed.

The second important step ahead has been made by Richmond and Horowitz in [7] and then by Parker and Richmond in [14] where the classical, Gaussian-based, SB formula has been extended to estimation problems under model misspecification, i.e. when the assumed parametric Gaussian model, say $\mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}))$, could differ from the true (possibly non parametric) one, indicated as $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. In other words, we allow the assumed parametric mean value $\boldsymbol{\gamma}(\boldsymbol{\theta})$ and the assumed parametric covariance matrix $\boldsymbol{\Pi}(\boldsymbol{\theta})$ to differ from the true $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ for every possible value of the parameter vector $\boldsymbol{\theta} \in \Theta$, i.e. $\mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta})) \neq \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \forall \boldsymbol{\theta} \in \Theta$. It is worth to underline that in the estimation framework under model misspecification, the FIM loses its classical statistical sense and it has to be substituted by the matrices $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{B}(\boldsymbol{\theta})$ defined in [8], Eqs. (1) and (7), respectively (see also [6,7]). Consequently, in this context, SB-type formulas could be exploited to obtain $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{B}(\boldsymbol{\theta})$ needed to evaluate the counterpart of the CRB in the presence of model misspecification, i.e. the Misspecified CRB (MCRB) [4,6–8,11]. In particular, in [7] the authors derived SB-type formulas for the "decoupled" scenario in which the unknown parameter vector $\boldsymbol{\theta} \in \Theta$ can be partitioned in two sub-vectors $\boldsymbol{\eta}$ and $\boldsymbol{\nu}$, named "deterministic" and "stochastic" parameter subvectors respectively, such that $\boldsymbol{\theta} = [\boldsymbol{\eta}^T, \boldsymbol{\nu}^T]^T$ and $\mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta})) \triangleq \mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\eta}), \boldsymbol{\Pi}(\boldsymbol{\nu})) \neq \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \forall \boldsymbol{\theta} \in \Theta$. The findings presented in [7] have been extended in [14] to include the coupling of deterministic and stochastic parameters. More formally, in [14], SB-type formulas have been derived for the following misspecified scenario $\mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta})) \triangleq \mathcal{CN}(\boldsymbol{\gamma}(\boldsymbol{\eta}, \boldsymbol{\omega}), \boldsymbol{\Pi}(\boldsymbol{\nu}, \boldsymbol{\omega})) \neq \mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \forall \boldsymbol{\theta} \in \Theta$ where the unknown parameter vector $\boldsymbol{\theta} \in \Theta$ is partitioned as $\boldsymbol{\theta} = [\boldsymbol{\eta}^T, \boldsymbol{\nu}^T, \boldsymbol{\omega}^T]^T$.

The natural extension of the works of Besson and Abramovich [18], Richmond and Horowitz [7] and Parker and Richmond [14] would be to derive SB-type formulas for parametric estimation problems involving CES distributed data under model misspecification. This paper aims exactly at filling this gap and obtaining some general "misspecified" SB formulas for CES distributed data.

Remark: Throughout this paper, we consider only the case of *real* parameter vectors. This is not a limitation, since we can always maps a complex vector in a real one simply by stacking its real and the imaginary parts. Clearly, the proposed derivation of the SB-type formulas could also be developed directly in the complex field by means of the Wirtinger calculus as in [7,19].

Notation: Throughout this paper, italics indicates scalar quantities (a, A), lower case and upper case boldface indicate column vectors ($\mathbf{a}$) and matrices ($\mathbf{A}$) respectively. Each entry of a matrix $\mathbf{A}$ is indicated as $a_{i,j} \triangleq [\mathbf{A}]_{i,j}$. * indicates the complex conjugation. The superscripts T and H indicates the transpose and the Hermitian operators, then $\mathbf{A}^H = (\mathbf{A}^*)^T$. Let $f(t)$ be a real scalar function, than $f'(t) \triangleq df(t)/dt$. Let $\mathbf{A}(\boldsymbol{\theta})$ be a matrix (or possibly vector or even scalar) function of the vector $\boldsymbol{\theta}$, then $\mathbf{A}_0 \triangleq \mathbf{A}(\boldsymbol{\theta}_0)$ while $\mathbf{A}_i^0 \triangleq \frac{\partial \mathbf{A}(\boldsymbol{\theta})}{\partial \theta_i}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$ and $\mathbf{A}_{ij}^0 \triangleq \frac{\partial^2 \mathbf{A}(\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$, where the vector $\boldsymbol{\theta}_0$ will be always explicitly defined in the paper. For two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semi-definite. Finally, for random variables or vectors, the notation $=_d$ stands for "has the same distribution as".

2. Problem setup

Let $\mathbf{x} = \{\mathbf{x}_l\}_{l=1}^{L}$, with $\mathbf{x}_l \in \mathbb{C}^N$, be a set of L independent complex random vectors (or *snapshots*) representing the available observations. We assume that each snapshot is sampled from a CES distribution [2,3], i.e., $\mathbf{x}_l \sim CES_N(\boldsymbol{\mu}_l,\ \boldsymbol{\Sigma},\ g)$, then its pdf can be expressed as:

$$p_X(\mathbf{x}_l) \triangleq p_X(\mathbf{x}_l; \boldsymbol{\mu}_l, \boldsymbol{\Sigma}) = c_{N,g} |\boldsymbol{\Sigma}|^{-1} g((\mathbf{x}_l - \boldsymbol{\mu}_l)^H \boldsymbol{\Sigma}^{-1} (\mathbf{x}_l - \boldsymbol{\mu}_l)) \quad (1)$$

where $c_{N,g}$ is a normalizing constant, $g(t) : \mathbb{R}^+ \to \mathbb{R}$ is the *density generator*, $\boldsymbol{\mu}_l = E_p\{\mathbf{x}_l\}$ is the mean value and $\boldsymbol{\Sigma}$ is a positive definite Hermitian matrix called *scatter matrix*. In the rest of this paper, we always assume that the scatter matrix $\boldsymbol{\Sigma}$ is of full rank, i.e. $\text{rank}(\boldsymbol{\Sigma}) = N$. From the Stochastic Representation Theorem [2], a CES distributed random vector can be expressed as:

$$\mathbf{x}_l =_d \boldsymbol{\mu}_l + \mathcal{R}\mathbf{T}\mathbf{u}_l, \quad (2)$$

where:

- $\mathbf{u}_l \sim U(\mathbb{C}S^N)$ is a N-dimensional complex random vector uniformly distributed on the unit hyper-sphere with $N-1$ topological dimension. As reported in [2] (Lemma 1), $E_p\{\mathbf{u}_l\} = 0$ and $E_p\{\mathbf{u}_l \mathbf{u}_l^H\} = (1/N)\mathbf{I}$ where $\mathbf{I}$ is the identity matrix of a suitable dimension.
- $\mathcal{R} \triangleq \sqrt{\mathcal{Q}}$ is a real and non-negative random variable called *modular variate*, while $\mathcal{Q}$ is called *second order modular variate*. Moreover, under the assumption that $\text{rank}(\boldsymbol{\Sigma}) = N$, we have that:

$$Q_l \triangleq (\mathbf{x}_l - \boldsymbol{\mu}_l)^H \boldsymbol{\Sigma}^{-1} (\mathbf{x}_l - \boldsymbol{\mu}_l) =_d \mathcal{Q}, \forall l \in \mathbb{N}. \quad (3)$$

As shown in [2], the pdf of $\mathcal{Q}$ has a one-to-one relation with density generator:

$$p_{\mathcal{Q}}(t) = \delta_{N,g}^{-1} t^{N-1} g(t), \quad (4)$$

where $\delta_{N,g} \triangleq \int_0^\infty t^{N-1} g(t) dt < \infty$. As a consequence of (3) and (4), the expectation of functions of the quadratic form Q_l, say

$h(Q_l)$, can be explicitly derived as:

$$E_{\mathcal{Q}}\{h(Q_l)\} \triangleq \int_0^{\infty} h(t)\, p_{\mathcal{Q}}(t)dt = \delta_{N,g}^{-1} \int_0^{\infty} h(t) t^{N-1} g(t)dt, \forall l \in \mathbb{N}. \tag{5}$$

It is clear from (5) that such expectation does not depend on the index l, since the pdf in (4) of the quadratic form Q_l is invariant with respect to (w.r.t.) l. For this reason, to avoid confusion, in the rest of the paper we always indicate $E_{\mathcal{Q}}\{h(Q_l)\}$ simply as $E_{\mathcal{Q}}\{h(Q)\}$.

- $\mathbf{T}$ is a complex $N \times N$ matrix with $\text{rank}(\mathbf{T}) = N$, such that $\boldsymbol{\Sigma} = \mathbf{T}\mathbf{T}^H$.
- If $E_{\mathcal{Q}}\{Q\} < \infty$ and $\text{rank}(\boldsymbol{\Sigma}) = N$, then the covariance matrix $\mathbf{M} = E_p\{(\mathbf{x}_l - \boldsymbol{\mu}_l)(\mathbf{x}_l - \boldsymbol{\mu}_l)^H\}$ of $\mathbf{x}_l$ can be decomposed as $\mathbf{M} = \sigma^2 \boldsymbol{\Sigma}$, where (see Theorem 4 in [2]):

$$\sigma^2 \triangleq \frac{E_{\mathcal{Q}}\{Q\}}{N}, \tag{6}$$

and σ^2 can be interpreted as the *statistical power* of the CES-distributed vector $\mathbf{x}_l$.

From the Stochastic Representation Theorem, it is clear that the representation of a CES distributed vector $\mathbf{x}_l$ is not uniquely determined by (2). In fact, $\mathbf{x}_l =_d \boldsymbol{\mu}_l + \mathcal{R}\mathbf{T}\mathbf{u}_l =_d \boldsymbol{\mu}_l + (c^{-1}\mathcal{R})(c\mathbf{T})\mathbf{u}_l, \forall c > 0$. From a different, yet equivalent, standpoint, this identifiability problem can be understood as an implicit consequence of the functional expression of a CES distribution. It is immediate to verify from (1) that $CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g(t)) \equiv CES_N(\boldsymbol{\mu}_l, c^2\boldsymbol{\Sigma}, g(t/c^2)), \forall c > 0$. As amply discussed in [2], this identifiability issue can be solved by posing a constraint on the modular variate $\mathcal{R}$ (and consequently, through (4), on the density generator $g(t)$), or on the scatter matrix $\boldsymbol{\Sigma}$. For convenience, we choose to put a constraint on $\mathcal{R}^2 = \mathcal{Q}$. Specifically in the rest of this paper, we always assume that $E_{\mathcal{Q}}\{Q\} = N$ and consequently, from (6), $\mathbf{M} = \boldsymbol{\Sigma}$, i.e. the scatter matrix equates the covariance matrix.

Due to the independence assumption, the joint pdf of the set $\mathbf{x}$ is given by the product of the marginal pdfs of each snapshot $\mathbf{x}_l$ given in (1):

$$\begin{aligned} p_X(\mathbf{x}) &\triangleq p_X(\mathbf{x}_1, \ldots, \mathbf{x}_L; \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_L, \boldsymbol{\Sigma}) \\ &= (c_{N,g})^L |\boldsymbol{\Sigma}|^{-L} \prod_{l=1}^{L} g((\mathbf{x}_l - \boldsymbol{\mu}_l)^H \boldsymbol{\Sigma}^{-1} (\mathbf{x}_l - \boldsymbol{\mu}_l)) \end{aligned} \tag{7}$$

As discussed in the Introduction, the pdf in (7) is a general and powerful model that it is able to statistically characterize the impulsive, heavy-tailed data behaviour in a variety of applications and allow us to overtake the Gaussianity assumption. Let us now focus on the clearing of the matched model assumption. Following the recent developments on this field, in this paper we consider the following mismatched situation:

- The acquired dataset $\mathbf{x} = \{\mathbf{x}_l\}_{l=1}^L$ is characterized by the *true* but *unknown* joint pdf given in (7). In particular, each data snapshot $\mathbf{x}_l$ follows a CES distribution with mean value $\boldsymbol{\mu}_l$, scatter matrix $\boldsymbol{\Sigma}$ and density generator $g(t)$, i.e. $\mathbf{x}_l \sim CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g), l = 1, \ldots, L$.
- In order to derive an inference algorithm, we assume that each snapshot of the dataset $\mathbf{x}$ is sampled from a CES distribution with a density generator $w(t)$, possibly different from $g(t)$ for all $t \in \mathbb{R}^+$, and a mean value $\boldsymbol{\gamma}_l(\boldsymbol{\theta})$ and a scatter matrix $\boldsymbol{\Pi}(\boldsymbol{\theta})$ parametrized by a deterministic parameter vector $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$ to be estimated. In particular, we allow the assumed marginal model $CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$ to differ from the true one, $CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g)$ for every $\boldsymbol{\theta} \in \Theta$.

This is a recurring scenario in array processing applications, where the mean value and/or the scatter matrix of the acquired snapshot vectors are assumed to be parametrized by a deterministic parameter vector whose components represent the Doppler frequency, the Direction of Arrivals (DOAs) of potential sources and so on.

The *assumed* marginal pdf of each snapshot $\mathbf{x}_l$ can then be expressed as:

$$\begin{aligned} f_X(\mathbf{x}_l; \boldsymbol{\theta}) &\triangleq f_X(\mathbf{x}_l; \boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta})) \\ &= c_{N,w} |\boldsymbol{\Pi}(\boldsymbol{\theta})|^{-1} w((\mathbf{x}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}))^H \boldsymbol{\Pi}(\boldsymbol{\theta})^{-1} (\mathbf{x}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}))) \end{aligned} \tag{8}$$

and, by exploiting the independence assumption, the joint pdf of $\mathbf{x} = \{\mathbf{x}_l\}_{l=1}^L$ can be obtained as:

$$\begin{aligned} f_X(\mathbf{x}; \boldsymbol{\theta}) &\triangleq f_X(\mathbf{x}_1, \ldots, \mathbf{x}_L; \boldsymbol{\gamma}_l(\boldsymbol{\theta}), \ldots, \boldsymbol{\gamma}_L(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta})) \\ &= (c_{N,w})^L |\boldsymbol{\Pi}(\boldsymbol{\theta})|^{-L} \prod_{l=1}^{L} w\Big((\mathbf{x}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}))^H \boldsymbol{\Pi}(\boldsymbol{\theta})^{-1} (\mathbf{x}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}))\Big) \end{aligned} \tag{9}$$

This scenario clearly represents an estimation problem in non-Gaussian data and in the presence of model misspecification. Let $\hat{\boldsymbol{\theta}}_f \triangleq \hat{\boldsymbol{\theta}}_f(\mathbf{x})$ be a, possibly mismatched, estimator of the parameter vector $\boldsymbol{\theta} \in \Theta$ derived under the assumed model $f_X(\mathbf{x}; \boldsymbol{\theta})$ in (9) while the data are characterized by the true model $p_X(\mathbf{x})$ in (7). Then, as discussed in [6–8], under suitable regularity conditions, a lower bound on the error covariance matrix of any *misspecified (MS)*-unbiased (see [6,8]) mismatched estimator $\hat{\boldsymbol{\theta}}_f$ exists and it is given by the MCRB defined as:

$$\mathbf{C}_p\left(\hat{\boldsymbol{\theta}}_f, \boldsymbol{\theta}_0\right) \triangleq E_p\left\{(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_f - \boldsymbol{\theta}_0)^T\right\} \geq \frac{1}{L}\mathbf{A}^{-1}(\boldsymbol{\theta}_0)\mathbf{B}(\boldsymbol{\theta}_0)\mathbf{A}(\boldsymbol{\theta}_0)^{-1} \tag{10}$$

where $\boldsymbol{\theta}_0$ is the so-called *pseudo-true parameter vector* defined in [4–6] as the parameter vector that minimize the Kullback–Leibler divergence (KLD) between the true and the assumed models:

$$\boldsymbol{\theta}_0 \triangleq \arg\min_{\boldsymbol{\theta} \in \Theta} \{D(p_X \| f_{\boldsymbol{\theta}})\} = \arg\min_{\boldsymbol{\theta} \in \Theta} \left\{-E_p\{\ln f_X(\mathbf{x}_l; \boldsymbol{\theta})\}\right\}, \tag{11}$$

where $D(p_X \| f_{\boldsymbol{\theta}}) \triangleq E_p\{\ln(p_X(\mathbf{x}_l)/f_X(\mathbf{x}_l; \boldsymbol{\theta}))\}$. The matrices $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ are defined as:

$$\left[\mathbf{A}(\boldsymbol{\theta}_0)\right]_{i,j} \triangleq \left[E_p\left\{\nabla_{\boldsymbol{\theta}}^T \nabla_{\boldsymbol{\theta}} \ln f_X(\mathbf{x}_l; \boldsymbol{\theta}_0)\right\}\right]_{i,j} = E_p\left\{\frac{\partial^2 \ln f_X(\mathbf{x}_l; \boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right\} \tag{12}$$

and

$$\begin{aligned} \left[\mathbf{B}(\boldsymbol{\theta}_0)\right]_{i,j} &\triangleq \left[E_p\left\{\nabla_{\boldsymbol{\theta}} \ln f_X(\mathbf{x}_l; \boldsymbol{\theta}_0) \nabla_{\boldsymbol{\theta}}^T \ln f_X(\mathbf{x}_l; \boldsymbol{\theta}_0)\right\}\right]_{i,j} \\ &= E_p\left\{\frac{\partial \ln f_X(\mathbf{x}_l; \boldsymbol{\theta})}{\partial\theta_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \cdot \frac{\partial \ln f_X(\mathbf{x}_l; \boldsymbol{\theta})}{\partial\theta_j}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right\}. \end{aligned} \tag{13}$$

For a deep theoretical analysis of the existence, the properties and the practical applicability of the MCRB, we refer the reader to [6–9] and to the references therein. The goal of this paper is to provide a general closed-form expressions of $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ in the aforementioned context. In other words, we aim at deriving two SB-type formulas for the evaluation of the matrices $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ that represent a generalization of the classical FIM for estimation problem under model misspecification. It is important to note that in the following derivations we always assume the existence and the uniqueness of the pseudo-true parameter vector $\boldsymbol{\theta}_0$ defined in (11). As it is clear from (10) and as it is amply discussed in [6–8], finding the $\boldsymbol{\theta}_0$ that minimizes the KLD between the true and the assumed model is a prerequisite for the evaluation of the MCRB since all the derivatives involved in the two matrices $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ have to be evaluated at $\boldsymbol{\theta}_0$. Finally, it is worth noticing that, under suitable regularity conditions on the true and the assumed pdfs, the definition of the pseudo-true parameter vector $\boldsymbol{\theta}_0$

in (11) can be expressed in an equivalent form as:

$$\left.\frac{\partial D(p_X\|f_{\boldsymbol{\theta}})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = -E_p\left\{\left.\frac{\partial \ln f_X(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right\} = 0, \quad i=1,\ldots,d. \tag{14}$$

As we will see in the next section, this equality will be exploited to evaluate explicitly the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$. We note in passing the similarity between the definition on $\boldsymbol{\theta}_0$ given in (14) and the condition in Eq. (9) of [18]. This fact highlights the strong theoretical parallelism between the classical matched theory and the misspecified framework as detailed in [6–8].

3. Slepian-Bangs formulas under misspecified CES models

This Section focuses on the derivation of the SB formulas for $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ defined in (12) and (13), provided that there exists a unique pseudo-true parameter vector $\boldsymbol{\theta}_0$ satisfying (11). We start by evaluating explicit expressions for the following quantities:

$$V_{ij}(\boldsymbol{\theta}_0) = \left.\frac{\partial \ln f_X(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \cdot \left.\frac{\partial \ln f_X(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad i,j=1,\ldots,d, \tag{15}$$

and

$$H_{ij}(\boldsymbol{\theta}_0) = \left.\frac{\partial^2 \ln f_X(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}, \quad i,j=1,\ldots,d. \tag{16}$$

The entries of the matrices $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ can then be obtained by taking the expectation operator, w.r.t. the true distribution $p_X(\mathbf{x})$, of $H_{ij}(\boldsymbol{\theta}_0)$ and $V_{ij}(\boldsymbol{\theta}_0)$, i.e. $[\mathbf{A}(\boldsymbol{\theta}_0)]_{ij} = E_p\{H_{ij}(\boldsymbol{\theta}_0)\}$ and $[\mathbf{B}(\boldsymbol{\theta}_0)]_{ij} = E_p\{V_{ij}(\boldsymbol{\theta}_0)\}$. Unfortunately, this expectation can be evaluated in closed form only in two particular cases, as we will detail in Sections 3.2 and 3.3. In all the other cases, numerical integration techniques, e.g. the Monte Carlo integration, could be exploited.

3.1. Evaluation of $V_{ij}(\boldsymbol{\theta}_0)$ and $H_{ij}(\boldsymbol{\theta}_0)$ and of their expectation w.r.t. the true data distribution

According to the general mismatched estimation problem discussed in the previous section, we consider here the general case in which, for each available snapshot $\mathbf{x}_l$, a parametric CES model $f_X(\mathbf{x}_l;\boldsymbol{\theta}) = CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$ is assumed, while actually each observation vector is distributed according to a different, possibly non-parametric CES data model, i.e. $\mathbf{x}_l \sim p_X(\mathbf{x}_l) = CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g)$.

From (9), it is immediate to verify that:

$$\begin{aligned}\left.\frac{\partial \ln f_X(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} &= -L\left.\frac{\partial \ln|\boldsymbol{\Pi}(\boldsymbol{\theta})|}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ &\quad + \sum_{l=1}^{L}\phi(G_l(\boldsymbol{\theta}_0))\left.\frac{\partial G_l(\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \\ &= -L\mathrm{tr}(\mathbf{P}_i^0) + \sum_{l=1}^{L}\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\end{aligned} \tag{17}$$

where

$$\left.\frac{\partial \ln|\boldsymbol{\Pi}|}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = \mathrm{tr}\left(\boldsymbol{\Pi}_0^{-1}\boldsymbol{\Pi}_i^0\right) = \mathrm{tr}(\mathbf{P}_i^0), \tag{18}$$

with $\phi(t) = w'(t)/w(t)$ and

$$\mathbf{P}_i^0 \triangleq \boldsymbol{\Pi}_0^{-1/2}\boldsymbol{\Pi}_i^0\boldsymbol{\Pi}_0^{-1/2}, \tag{19}$$

$$G_l^0 \triangleq G_l(\boldsymbol{\theta}_0) \triangleq (\mathbf{x}_l - \boldsymbol{\gamma}_l^0)^H\boldsymbol{\Pi}_0^{-1}(\mathbf{x}_l - \boldsymbol{\gamma}_l^0), \tag{20}$$

where, for notational simplicity, $\boldsymbol{\gamma}_l^0 \triangleq \boldsymbol{\gamma}_l(\boldsymbol{\theta}_0)$ and $\boldsymbol{\Pi}_0 \triangleq \boldsymbol{\Pi}(\boldsymbol{\theta}_0)$. Then, an explicit expression of $V_{ij}(\boldsymbol{\theta}_0)$ in (15) is given by:

$$\begin{aligned}V_{ij}(\boldsymbol{\theta}_0) &= L^2\mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) - L\mathrm{tr}(\mathbf{P}_i^0)\sum_{l=1}^{L}\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_j} \\ &\quad - L\mathrm{tr}(\mathbf{P}_j^0)\sum_{l=1}^{L}\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i} \\ &\quad + \sum_{l=1}^{L}\sum_{m=1}^{L}\phi(G_l^0)\phi(G_m^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j},\end{aligned} \tag{21}$$

where:

$$\frac{\partial G_l^0}{\partial\theta_i} = -2\mathrm{Re}\left[(\mathbf{x}_l - \boldsymbol{\gamma}_l^0)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right] - (\mathbf{x}_l - \boldsymbol{\gamma}_l^0)^H\mathbf{S}_i^0(\mathbf{x}_l - \boldsymbol{\gamma}_l^0), \tag{22}$$

and

$$\mathbf{S}_i^0 = \boldsymbol{\Pi}_0^{-1}\boldsymbol{\Pi}_i^0\boldsymbol{\Pi}_0^{-1}. \tag{23}$$

The term $H_{ij}(\boldsymbol{\theta}_0)$ in (16) can be obtained, through direct calculation, from (17) as:

$$\begin{aligned}H_{ij}(\boldsymbol{\theta}_0) &= L\mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0 - \mathbf{P}_{ij}^0) + \sum_{l=1}^{L}\phi'(G_l^0)\frac{\partial G_l^0}{\partial\theta_j}\frac{\partial G_l^0}{\partial\theta_i} \\ &\quad + \sum_{l=1}^{L}\phi(G_l^0)\frac{\partial^2 G_l^0}{\partial\theta_i\partial\theta_j}\end{aligned} \tag{24}$$

where:

$$\mathbf{P}_{ij}^0 \triangleq \boldsymbol{\Pi}_0^{-1/2}\boldsymbol{\Pi}_{ij}^0\boldsymbol{\Pi}_0^{-1/2} \tag{25}$$

and $\phi'(t) = w''(t)/w(t) - (w'(t))^2/w^2(t)$. After having obtained the terms $V_{ij}(\boldsymbol{\theta}_0)$ and $H_{ij}(\boldsymbol{\theta}_0)$, we have to evaluate their expectations w.r.t. the true distribution $p_X(\mathbf{x})$. Since all the derivatives in (21) and (24) have to be evaluated in the pseudo-true parameter vector $\boldsymbol{\theta}_0$, we can exploit the equality established in (14). Specifically, from (17), we have that:

$$\left.\frac{\partial D(p_X\|f_{\boldsymbol{\theta}})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} = L\mathrm{tr}(\mathbf{P}_i^0) - \sum_{l=1}^{L}E_p\left\{\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\right\} = 0, \quad i=1,\ldots,d, \tag{26}$$

and consequently,

$$\sum_{l=1}^{L}E_p\left\{\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\right\} = L\mathrm{tr}(\mathbf{P}_i^0), \quad i=1,\ldots,d. \tag{27}$$

Now, taking the expectation operator w.r.t. $p_X(\mathbf{x})$ of the term $V_{ij}(\boldsymbol{\theta}_0)$ in (21) and by exploiting the equality in (27), the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$ can be expressed as:

$$[\mathbf{B}(\boldsymbol{\theta}_0)]_{ij} = -L^2\mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) + \sum_{l=1}^{L}\sum_{m=1}^{L}E_p\left\{\phi(G_l^0)\phi(G_m^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\}. \tag{28}$$

Similarly, the matrix $\mathbf{A}(\boldsymbol{\theta}_0)$ can be obtained by taking the expectation operator w.r.t. $p_X(\mathbf{x})$ of the term $H_{ij}(\boldsymbol{\theta}_0)$ in (24) as:

$$\begin{aligned}[\mathbf{A}(\boldsymbol{\theta}_0)]_{ij} &= L\mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0 - \mathbf{P}_{ij}^0) + \sum_{l=1}^{L}E_p\left\{\phi'(G_l^0)\frac{\partial G_l^0}{\partial\theta_j}\frac{\partial G_l^0}{\partial\theta_i}\right\} \\ &\quad + \sum_{l=1}^{L}E_p\left\{\phi(G_l^0)\frac{\partial^2 G_l^0}{\partial\theta_i\partial\theta_j}\right\}.\end{aligned} \tag{29}$$

As we can see from (21) and (24), the expressions of $V_{ij}(\boldsymbol{\theta}_0)$ and $H_{ij}(\boldsymbol{\theta}_0)$ are highly involved from an analytical standpoint, and consequently it is impossible to derive in closed form their expectations in the general case. As discussed in Appendix A, closed form expressions can be obtained when the random terms in (21) and (24) satisfy certain independence conditions. Fortunately, there are two scenarios of great practical interest in which such conditions are met and consequently a closed form expression for $\mathbf{A}(\boldsymbol{\theta}_0)$ and $\mathbf{B}(\boldsymbol{\theta}_0)$ can be obtained. These special cases are:

Scenario 1: The true marginal pdf is an unspecified CES model, i.e. $p_X(\mathbf{x}_l) = CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g)$, while the assumed pdf is a parametric complex Gaussian model, i.e. $f_X(\mathbf{x}_l;\boldsymbol{\theta}) = \mathcal{CN}(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}))$.

Scenario 2: The Scenario 2 is characterized by two assumptions:

A1: The true and the assumed pdfs share the same parametric mean value $\boldsymbol{\gamma}_l(\boldsymbol{\theta})$ and the same parametric scatter matrix $\boldsymbol{\Pi}(\boldsymbol{\theta})$ while the misspecification is caused by a wrong assumption on the density generator $w(t)$. More formally, posit that the true marginal pdf of each snapshot is given by a parametric CES distribution such that $p_X(\mathbf{x}_l) \triangleq p_X(\mathbf{x}_l;\bar{\boldsymbol{\theta}}) = CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g)$, where $\boldsymbol{\mu}_l = \boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}})$ and $\boldsymbol{\Sigma} = \boldsymbol{\Pi}(\bar{\boldsymbol{\theta}})$ for a given $\bar{\boldsymbol{\theta}} \in \Theta$. The assumed model is instead another parametric CES distribution that share the same parameterization of the true one but may have a different density generator, i.e. $f_X(\mathbf{x}_l;\boldsymbol{\theta}) = CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$, $\boldsymbol{\theta} \in \Theta$, and $g(t) \neq w(t), \forall t \in \mathbb{R}^+$. Note that this scenario is a particular case of the more general class of misspecified problems discussed in [8], Section II.D.

A2: The true parameter vector $\bar{\boldsymbol{\theta}}$ and the pseudo-true parameter vector $\boldsymbol{\theta}_0$ are equals. In particular, $\bar{\boldsymbol{\theta}}$ is the solution of the optimization problem in (11).

Note that these two assumption are verified for the scatter matrix estimation problem discussed in [8].

The Scenario 1 describes a common practice in array processing applications. In fact, when the true data model is unknown, a prevalent choice is to assume a simple Gaussian model that guarantees an easy derivation and a consequent real-time implementation of the estimation algorithm.

The Scenario 2 is a bit different, since it imply the a-priori knowledge of the functional form of the true parametric mean value $\boldsymbol{\gamma}_l(\boldsymbol{\theta})$ and of the parametric scatter matrix $\boldsymbol{\Pi}(\boldsymbol{\theta})$. There are, however, a variety of practical applications in which this a-priori information is indeed available to the user. As an example, one can think of array signal processing applications in which the a-priori knowledge of the array geometry leads to a correct specification of the parametrized mean value and covariance matrix of the collected snapshots, while the uncertainty on the statistical disturbance model could cause a wrong choice of the density generator. Of course, there are cases in which also the knowledge of the array geometry could be incorrect or partial, and consequently the assumed mean value and scatter matrix differ from the true ones (see [7] for more details).

3.2. Scenario 1

In this subsection we provide the SB formulas, i.e. the closed form expressions of the matrices $\mathbf{A}(\boldsymbol{\theta}_0)$ in (12) and $\mathbf{B}(\boldsymbol{\theta}_0)$ in (13), for the Scenario 1. We start by noticing that the assumed complex Gaussian model belongs to the CES class, i.e. $f_X(\mathbf{x}_l;\boldsymbol{\theta}) = CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$, where $w(t) = \exp(-t)$. Consequently, it is immediate to verify that $\phi(t) = -1$ and $\phi'(t) = 0$. The matrix $\mathbf{B}(\boldsymbol{\theta}_0)$ can be evaluated from (28) by using the fact that $\phi(G_l) = -1$:

$$[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} = -L^2 \mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) + \sum_{l=1}^{L}\sum_{m=1}^{L} E_p\left\{\frac{\partial G_l^0}{\partial \theta_i}\frac{\partial G_m^0}{\partial \theta_j}\right\}. \tag{30}$$

Following the procedure discussed in Appendix B, the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$ can be expressed as:

$$\begin{aligned}[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} &= 2\sum_{l=1}^{L}\mathrm{Re}\Bigg[\left(\boldsymbol{\Pi}_i^0\boldsymbol{\Pi}_0^{-1}\mathbf{r}_l^0 + \frac{\partial \boldsymbol{\gamma}_l^0}{\partial \theta_i}\right)^H \boldsymbol{\Pi}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{\Pi}_0^{-1} \\ &\times\left(\boldsymbol{\Pi}_j^0\boldsymbol{\Pi}_0^{-1}\mathbf{r}_l^0 + \frac{\partial \boldsymbol{\gamma}_l^0}{\partial \theta_j}\right)\Bigg] + L\left(\frac{E_Q\{Q^2\}}{N(N+1)} - 1\right)\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma})\mathrm{tr}(\mathbf{S}_j^0\boldsymbol{\Sigma}) \\ &+ L\frac{E_Q\{Q^2\}}{N(N+1)}\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}\mathbf{S}_j^0\boldsymbol{\Sigma}),\end{aligned} \tag{31}$$

where:

$$\mathbf{r}_l^0 \triangleq \boldsymbol{\mu}_l - \boldsymbol{\gamma}_l^0. \tag{32}$$

Note that all the derivatives have to be evaluated in the pseudo-true parameter vector $\boldsymbol{\theta}_0$ defined in (11).

Let us evaluate the matrix $\mathbf{A}(\boldsymbol{\theta}_0)$. From (24) and $\phi(G_l) = -1$, and $\phi'(G_l) = 0$, through direct calculation (see also Appendix B) we obtain:

$$\begin{aligned}\alpha_l^{ij}(\boldsymbol{\theta}_0) \triangleq E_p\left\{\frac{\partial^2 G_l^0}{\partial\theta_i\partial\theta_j}\right\} &= 2\mathrm{Re}\left[\frac{\partial \boldsymbol{\gamma}_l^{0\,H}}{\partial\theta_i}\boldsymbol{\Pi}_0^{-1}\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_j}\right] \\ &+ 2\mathrm{Re}\left[\mathbf{r}_l^{0H}\left(\mathbf{S}_j^0\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_i} + \mathbf{S}_i^0\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_j} - \boldsymbol{\Pi}_0^{-1}\frac{\partial^2 \boldsymbol{\gamma}_l^0}{\partial\theta_i\partial\theta_j}\right)\right] \\ &+ \mathbf{r}_l^{0H}\boldsymbol{\Pi}_0^{-1/2}(\mathbf{P}_i^0\mathbf{P}_j^0 + \mathbf{P}_j^0\mathbf{P}_i^0 - \mathbf{P}_{ij}^0)\boldsymbol{\Pi}_0^{-1/2}\mathbf{r}_l^0 \\ &+ \mathrm{tr}\big[(\mathbf{P}_i^0\mathbf{P}_j^0 + \mathbf{P}_j^0\mathbf{P}_i^0 - \mathbf{P}_{ij}^0)\boldsymbol{\Pi}_0^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Pi}_0^{-1/2}\big].\end{aligned} \tag{33}$$

Moreover, as in (31), we used constraint on Q_l. Finally, by inserting (33) in (24), we obtain:

$$[\mathbf{A}(\boldsymbol{\theta}_0)]_{i,j} = L\mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0 - \mathbf{P}_{ij}^0) - \sum_{l=1}^{L}\alpha_l^{ij}(\boldsymbol{\theta}_0). \tag{34}$$

The expressions (31) and (34) represent the SB formulas for the mismatched Scenario 1.

3.3. Scenario 2

This Subsection focuses on the Scenario 2, i.e. the case in which the true and the assumed pdfs are CES distributions that share the same parametrized mean value and scatter matrix but are characterized by different density generators. From the proof provided in Appendix C, the matrices $\mathbf{B}(\boldsymbol{\theta}_0)$ and $\mathbf{A}(\boldsymbol{\theta}_0)$ can be expressed respectively as:

$$\begin{aligned}[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} = [\mathbf{B}(\bar{\boldsymbol{\theta}})]_{i,j} &= \frac{2}{N}E_Q\{Q\phi^2(Q)\}\sum_{l=1}^{L}\mathrm{Re}\left[\left(\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_j}\right] \\ &+ L\left(\frac{E_Q\{Q^2\phi^2(Q)\}}{N(N+1)} - 1\right)\mathrm{tr}(\mathbf{P}_i)\mathrm{tr}(\mathbf{P}_j) + L\frac{E_Q\{Q^2\phi^2(Q)\}}{N(N+1)}\mathrm{tr}(\mathbf{P}_i\mathbf{P}_j),\end{aligned} \tag{35}$$

$$\begin{aligned}[\mathbf{A}(\boldsymbol{\theta}_0)]_{i,j} = [\mathbf{A}(\bar{\boldsymbol{\theta}})]_{i,j} &= \frac{2}{N}\big(E_Q\{Q\phi'(Q)\} + NE_Q\{\phi(Q)\}\big)\sum_{l=1}^{L}\mathrm{Re} \\ &\times\left[\left(\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial \boldsymbol{\gamma}_l^0}{\partial\theta_j}\right] + L\left(\frac{E_Q\{Q^2\phi'(Q)\}}{N(N+1)} - 1\right)\mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0) \\ &+ L\frac{E_Q\{Q^2\phi'(Q)\}}{N(N+1)}\mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0).\end{aligned} \tag{36}$$

The expressions (35) and (36) represent the SB formulas for the mismatched Scenario 2. It is worth to recall that the previous two formulas can be applied only if, in the particular estimation problem an hand, the pseudo-true parameter vector equates the true parameter vector, i.e. when $\boldsymbol{\theta}_0 = \bar{\boldsymbol{\theta}}$. The reason for this restriction will be clarified in Appendix C.

4. Relationship to previous results

The aim of this Section is to show that the SB formulas derived for the Scenarios 1 and 2 encompass all the previously derived SB formulas as special cases. In particular, both the SB formula for CES distributions under perfect model specification, proposed in [18], and the SB formulas for the scatter/covariance matrix estimation of CES distributed vectors under misspecification of the density generator, proposed in [8,11], can be obtained as special cases of the SB formulas shown here for the Scenario 2. In addition, the SB formulas for misspecified Gaussian models, proposed in [7], are special cases of the SB formulas discussed here for the Scenario 1. We note, in passing, that the SB-type formulas derived in [14] can also be obtained as a particular case of the ones proposed in this paper in (31) and (34). This can be easily done by partitioning the unknown parameter vector as $\boldsymbol{\theta} = [\boldsymbol{\eta}^T, \boldsymbol{\nu}^T, \boldsymbol{\omega}^T]^T$ and by taking into account the particular parameterization of the assumed mean value and of the assumed scatter matrix supposed in [14], i.e. $\boldsymbol{\gamma}(\boldsymbol{\theta}) = \boldsymbol{\gamma}(\boldsymbol{\eta}, \boldsymbol{\omega})$ and $\boldsymbol{\Pi}(\boldsymbol{\theta}) = \boldsymbol{\Pi}(\boldsymbol{\nu}, \boldsymbol{\omega})$.

4.1. The SB formula under correctly specified CES models [18]

The SB formula for correctly specified CES model has been derived in [18]. Using the formalism introduced in this paper, two parametric CES models are said to be correctly specified if there exists a vector $\bar{\boldsymbol{\theta}} \in \Theta$, such that the assumed CES distribution in (9) equates the true CES distribution in (7). More formally, the CES model $f_X(\mathbf{x}; \boldsymbol{\theta})$ is said to be correctly specified if there exists $\bar{\boldsymbol{\theta}} \in \Theta$, such that $f_X(\mathbf{x}; \bar{\boldsymbol{\theta}}) = p_X(\mathbf{x})$ and, in particular, $\boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}}) = \boldsymbol{\mu}_l$, $\boldsymbol{\Pi}(\bar{\boldsymbol{\theta}}) = \boldsymbol{\Sigma}$ and $g(t) \equiv w(t)$. As proved in [6], under correctly specified model, we have that $\bar{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$ and $\mathbf{B}(\bar{\boldsymbol{\theta}}) = -\mathbf{A}(\bar{\boldsymbol{\theta}})$, where $\mathbf{B}(\bar{\boldsymbol{\theta}})$ is the classical FIM evaluated at the true parameter vector $\bar{\boldsymbol{\theta}}$.

In the following, we will show that the SB formula derived in [18] can be considered as a special case of the one obtained in Section 3.3. In fact, according to the matched model assumption, we can define the true model as $p_X(\mathbf{x}) = f_X(\mathbf{x}; \bar{\boldsymbol{\theta}}) = CES_N(\boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}}), \boldsymbol{\Pi}(\bar{\boldsymbol{\theta}}), w)$, while the assumed parametric model is $f_X(\mathbf{x}; \boldsymbol{\theta}) = CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$ with $\boldsymbol{\theta} \in \Theta$. With this in mind, we can exploit the result in Eq. (11) of [18], that is $E_Q\{Q_l \bar{\phi}(Q_l)\} = -N$, where $\bar{\phi}(t) \triangleq g'(t)/g(t)$. Finally, as shown in [6], we have that $\mathbf{B}(\boldsymbol{\theta}_0) = \mathbf{B}(\bar{\boldsymbol{\theta}}) = -\mathbf{A}(\bar{\boldsymbol{\theta}}) = \mathrm{FIM}(\bar{\boldsymbol{\theta}})$. Then by exploiting (35), we obtain:

$$
\begin{aligned}
[\mathbf{B}(\bar{\boldsymbol{\theta}})]_{i,j} = [\mathrm{FIM}(\bar{\boldsymbol{\theta}})]_{i,j} &= \frac{2}{N} E_Q\{Q\bar{\phi}^2(Q)\} \sum_{l=1}^{L} \mathrm{Re}\left[\left(\frac{\partial \bar{\boldsymbol{\gamma}}_l}{\partial \theta_i} \right)^H \bar{\boldsymbol{\Pi}}^{-1} \frac{\partial \bar{\boldsymbol{\gamma}}_l}{\partial \theta_j} \right] \\
&+ L \left(\frac{E_Q\{Q^2\bar{\phi}^2(Q)\}}{N(N+1)} - 1 \right) \mathrm{tr}(\bar{\boldsymbol{\Pi}}^{-1}\bar{\boldsymbol{\Pi}}_i)\mathrm{tr}(\bar{\boldsymbol{\Pi}}^{-1}\bar{\boldsymbol{\Pi}}_j) \\
&+ L \frac{E_Q\{Q^2\bar{\phi}^2(Q)\}}{N(N+1)} \mathrm{tr}(\bar{\boldsymbol{\Pi}}^{-1}\bar{\boldsymbol{\Pi}}_i\bar{\boldsymbol{\Pi}}^{-1}\bar{\boldsymbol{\Pi}}_j),
\end{aligned} \tag{37}
$$

where $\bar{\boldsymbol{\Pi}} \triangleq \boldsymbol{\Pi}(\bar{\boldsymbol{\theta}})$ and $\bar{\boldsymbol{\Pi}}_i \triangleq \frac{\partial \boldsymbol{\Pi}(\boldsymbol{\theta})}{\partial \theta_i}|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}}$. It is immediate to verify that the matrix defined in (37) is exactly the same as the FIM given in Eq. (20) of [18].

4.2. The SB formulas for scatter matrix estimation under misspecification of the density generator [8,11]

In [8,11], SB formulas for the scatter matrix estimation in CES distributed vectors under misspecification of the density generator have been proposed. It is easy to verify that this scenario is a special case of the more general Scenario 2 discussed in Section 3.3. In particular, according to [8,11], the dataset $\mathbf{x}$ is considered to be composed of L independent, *zero mean* snapshots distributed as $\mathbf{x}_l \sim CES_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ where $\boldsymbol{\mu} \equiv \mathbf{0}$ and $\boldsymbol{\Sigma} = \boldsymbol{\Pi}(\boldsymbol{\theta})$, for a given $\bar{\boldsymbol{\theta}} \in \Theta$ that is the true parameter vector. For each snapshot, we assume a marginal pdf $f_X(\mathbf{x}_l; \boldsymbol{\theta}) = CES_N(\mathbf{0}, \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$, i.e. we misspecified the density generator. Since we aim at finding SB formulas for the estimation of the scatter matrix itself, we have that $\boldsymbol{\theta} \triangleq \mathrm{vecs}(\boldsymbol{\Pi}(\boldsymbol{\theta})) \equiv \mathrm{vecs}(\boldsymbol{\Pi})$, where vecs is the operator that maps a symmetric $N \times N$ matrix $\boldsymbol{\Pi}$ in a $N(N+1)/2$-dimensional vector whose entries are the elements of the upper (or lower) submatrix of $\boldsymbol{\Pi}$. A similar notation holds for the true parameter vector, and in particular $\bar{\boldsymbol{\theta}} = \mathrm{vecs}(\bar{\boldsymbol{\Pi}})$. Note that, in the following, we assume that the entries of $\boldsymbol{\theta}$ (and then of the scatter matrix) are real numbers. Finally, the results in [8,11] can be readily derived by (35) and (36) by posing $L = 1$, $\frac{\partial \boldsymbol{\gamma}_l^0}{\partial \theta_i} \equiv 0$ and $E_Q\{Q\phi(Q)\} = -N$ (see Appendix B):

$$
[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} = \frac{E_Q\{Q^2\phi^2(Q)\}}{N(N+1)} \left[\mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) + \mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0) \right] - \mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0), \tag{38}
$$

$$
[\mathbf{A}(\boldsymbol{\theta}_0)]_{i,j} = \frac{E_Q\{Q^2\phi'(Q)\}}{N(N+1)} \left[\mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0) + \mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) \right] - \mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0), \tag{39}
$$

where, in this case, the matrix $\mathbf{P}_i^0$ (or $\mathbf{P}_j^0$) can be expressed as $\mathbf{P}_i^0 = \boldsymbol{\Pi}_0^{-1}\mathbf{A}_i = \bar{\boldsymbol{\Pi}}^{-1}\mathbf{A}_i$ where $\boldsymbol{\Pi}_0$ is the matrix that minimizes the KLD between the true and the assumed distributions, and then, due to the Assumption A2, it is equal to the true covariance matrix, i.e. $\boldsymbol{\Pi}_0 = \bar{\boldsymbol{\Pi}} = \boldsymbol{\Pi}(\bar{\boldsymbol{\theta}}) = \boldsymbol{\Sigma}$. $\mathbf{A}_i$ is a matrix defined as $\mathbf{A}_i \triangleq \frac{\partial \boldsymbol{\Pi}(\boldsymbol{\theta})}{\partial \theta_i} \equiv \frac{\partial \boldsymbol{\Pi}}{\partial \theta_i}$.

4.3. The SB formulas for misspecified Gaussian models [7]

In this subsection, we briefly show how to obtain the SB formulas provided in [7] by using the general results discussed here for the Scenario 1. The SB formulas for misspecified Gaussian models can, in fact, be derived as a special case of the ones given in (31) and (34) by positing as true distribution the parametric complex Gaussian model, i.e. $p_X(\mathbf{x}_l) = CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, \exp(-t)) = \mathcal{CN}(\boldsymbol{\mu}_l, \boldsymbol{\Sigma})$, while the assumed marginal distribution is still given by $f_X(\mathbf{x}_l; \boldsymbol{\theta}) = CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), \exp(-t)) = \mathcal{CN}(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}))$. Note that the true and the assumed density generators are equal and it is simply given by $g(t) = w(t) = \exp(-t)$. This fact can be used to evaluate the term $E_Q\{Q^2\}$ as:

$$
\begin{aligned}
E_Q\{Q^2\} &= \int_0^\infty \exp(-t) t^{N+1} \delta_{N,g}^{-1} = -[\exp(-t) t^{N+1} \delta_{N,g}^{-1}]_0^\infty \\
&\quad + (N+1) \int_0^\infty t^N \exp(-t) \delta_{N,g}^{-1} dt.
\end{aligned} \tag{40}
$$

Since, $[\exp(-t) t^{N+1} \delta_{N,g}^{-1}]_0^\infty = 0$, we have that

$$
\begin{aligned}
E_Q\{Q^2\} &= (N+1) \int_0^\infty t^N \exp(-t) \delta_{N,g}^{-1} dt \\
&= -[\exp(-t) t^N \delta_{N,g}^{-1}]_0^\infty \\
&\quad + N(N+1) \int_0^\infty t^{N-1} \exp(-t) \delta_{N,g}^{-1} dt = N(N+1),
\end{aligned} \tag{41}
$$

in which we used the fact that $\int_0^\infty t^{N-1}\exp(-t)\delta_{N,g}^{-1}dt = 1$. Consequently, the entries of the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$ in (31) can be readily expressed as:

$$[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} = 2\sum_{l=1}^{L}\mathrm{Re}\Big[\Big(\boldsymbol{\Pi}_i^0\boldsymbol{\Pi}_0^{-1}\mathbf{r}_l^0 + \frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\Big)^H\boldsymbol{\Pi}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{\Pi}_0^{-1} \times\Big(\boldsymbol{\Pi}_j^0\boldsymbol{\Pi}_0^{-1}\mathbf{r}_l^0 + \frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_j}\Big)\Big] + L\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}\mathbf{S}_j^0\boldsymbol{\Sigma}), \tag{42}$$

where $\mathbf{r}_l^0 \triangleq \boldsymbol{\mu}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}_0)$.

4.3.1. The so-called Generalized Slepian formulas [7]: $\boldsymbol{\Pi}$ is constant w.r.t. $\boldsymbol{\theta}$

In this particular case, we have that all the derivatives of $\boldsymbol{\Pi}$ w.r.t. $\boldsymbol{\theta}$ are zero. In particular, $\mathbf{S}_i = \mathbf{S}_j = \mathbf{P}_i = \mathbf{P}_j = \mathbf{P}_{ij} = \mathbf{0}$. Then, by posing $L = 1$ as in [7], the entries of the matrices $\mathbf{B}(\boldsymbol{\theta}_0)$ and $\mathbf{A}(\boldsymbol{\theta}_0)$ can be obtained from (42) and (34) respectively, as:

$$[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} = 2\mathrm{Re}\left[\left(\frac{\partial\boldsymbol{\gamma}^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}^0}{\partial\theta_j}\right] \tag{43}$$

and

$$[\mathbf{A}(\boldsymbol{\theta}_0)]_{i,j} = -2\mathrm{Re}\left[\left(\frac{\partial\boldsymbol{\gamma}^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}^0}{\partial\theta_j}\right] - 2\mathrm{Re}\left[(\mathbf{r}^0)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial^2\boldsymbol{\gamma}^0}{\partial\theta_i\partial\theta_j}\right], \tag{44}$$

where $\mathbf{r}^0 \triangleq \boldsymbol{\mu} - \boldsymbol{\gamma}(\boldsymbol{\theta}_0)$. It is immediate to verify that these two expression are exactly the same of Eqs. (38) and (40) derived in [7].

4.3.2. The so-called Generalized Bangs formulas [7]: $\boldsymbol{\gamma}$ is constant w.r.t. $\boldsymbol{\theta}$

Here we suppose that the mean value $\boldsymbol{\gamma}$ is independent of the parameter vector $\boldsymbol{\theta}$, consequently we have that $\frac{\partial\boldsymbol{\gamma}}{\partial\theta_i} = 0$ and $\frac{\partial^2\boldsymbol{\gamma}}{\partial\theta_i\partial\theta_j} = 0$. Then, if $L = 1$, the entries of the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$ in (42) can be readily expressed as:

$$[\mathbf{B}(\boldsymbol{\theta}_0)]_{i,j} = 2\mathrm{Re}\Big[\big(\boldsymbol{\Pi}_i^0\boldsymbol{\Pi}_0^{-1}\mathbf{r}^0\big)^H\boldsymbol{\Pi}_0^{-1}\boldsymbol{\Sigma}\boldsymbol{\Pi}_0^{-1}\big(\boldsymbol{\Pi}_j^0\boldsymbol{\Pi}_0^{-1}\mathbf{r}_l^0\big)\Big] + \mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}\mathbf{S}_j^0\boldsymbol{\Sigma}), \tag{45}$$

from which one deduces straightforwardly Eq. (44) of [7] with $\mathbf{r}^0 \triangleq \boldsymbol{\mu} - \boldsymbol{\gamma}(\boldsymbol{\theta}_0)$. Finally, let us now derive the matrix $\mathbf{A}(\boldsymbol{\theta}_0)$ for the particular scenario at hand. From (34), we get:

$$[\mathbf{A}(\boldsymbol{\theta}_0)]_{i,j} = \mathrm{tr}(\mathbf{P}_i^0\mathbf{P}_j^0 - \mathbf{P}_{ij}^0) - \mathbf{r}_l^{0H}\boldsymbol{\Pi}_0^{-1/2}(\mathbf{P}_i^0\mathbf{P}_j^0 + \mathbf{P}_j^0\mathbf{P}_i^0 - \mathbf{P}_{ij}^0)\boldsymbol{\Pi}_0^{-1/2}\mathbf{r}_l^0 + \mathrm{tr}\big[(\mathbf{P}_i^0\mathbf{P}_j^0 + \mathbf{P}_j^0\mathbf{P}_i^0 - \mathbf{P}_{ij}^0)\boldsymbol{\Pi}_0^{-1/2}\boldsymbol{\Sigma}\boldsymbol{\Pi}_0^{-1/2}\big]. \tag{46}$$

which is the same as that given in Eq. (46) of [7].

5. Conclusion

The aim of this paper was to provide SB formulas for CES distributed data under model misspecification and thus to fill a theoretical and practical gap in the Signal Processing literature. Furthermore, we have shown that the proposed SB formulas encompass all the previously derived expressions as special cases. Moreover, these new SB formulas involved relatively slight modifications with respects to the classical counterpart obtained for Gaussian and correctly specified data models with only an expectation of some scalar functions to derive or to evaluate numerically. The practical importance of the proposed expressions is in the fact that allow us to easily evaluate Misspecified Cramér-Rao Bounds for a lot of applications that are characterized by a non-Gaussian and heavy-tailed data behaviour along with a model misspecification.

Future works will focus exactly on the application of the derived SB formulas to a plethora of engineering problems such as the Direction of Arrivals (DOA) estimation in array processing and the structured (i.e. Toeplitz) covariance/scatter matrix estimation for adaptive detection algorithm in non-Gaussian scenarios.

Acknowledgement

The work of Stefano Fortunati has been supported by the Air Force Office of Scientific Research under award number FA9550-17-1-0065.

The work of M. N. El Korso and A. Renaux has been supported by ANR-Astrid MARGARITA.

Appendix A. Some considerations on the expectation of the terms $V_{ij}(\boldsymbol{\theta}_0)$ and $H_{ij}(\boldsymbol{\theta}_0)$

The aim of this appendix is to show under which conditions the expectations $E_p\{V_{ij}(\boldsymbol{\theta}_0)\}$ and $E_p\{H_{ij}(\boldsymbol{\theta}_0)\}$ could be evaluated in closed form.

Keeping in mind the Stochastic Representation of a CES distributed random vector $\mathbf{x}_l$ given in (2), let us define the vector $\mathbf{z}_l = \mathbf{T}\mathbf{u}_l$. Then, by recalling that $\mathbf{r}_l^0 \triangleq \boldsymbol{\mu}_l - \boldsymbol{\gamma}_l^0$ (see (32)), the term $\frac{\partial G_l^0}{\partial\theta_i}$ in (22) can be rewritten as:

$$\frac{\partial G_l^0}{\partial\theta_i} = -\left(\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1}(Q_l^{1/2}\mathbf{z}_l + \mathbf{r}_l^0) - (Q_l^{1/2}\mathbf{z}_l + \mathbf{r}_l^0)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i} - (Q_l^{1/2}\mathbf{z}_l + \mathbf{r}_l^0)^H\mathbf{S}_i^0(Q_l^{1/2}\mathbf{z}_l + \mathbf{r}_l^0), \tag{A.1}$$

where $\mathbf{S}_i^0$ is defined in (23). Let us define the following vector and scalar quantities:

$$\mathbf{h}_{il} = \mathbf{S}_i^0\mathbf{r}_l^0 + \boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i} \tag{A.2}$$

and

$$A_{il} = (\mathbf{r}_l^0)^H\mathbf{S}_i^0\mathbf{r}_l^0 + (\mathbf{r}_l^0)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i} + \left(\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}^{-1}\mathbf{r}_l^0. \tag{A.3}$$

Then (A.1) becomes

$$\frac{\partial G_l^0}{\partial\theta_i} = -Q_l\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l - Q_l^{1/2}\mathbf{z}_l^H\mathbf{h}_{il} - Q_l^{1/2}\mathbf{h}_{il}^H\mathbf{z}_l - A_{il}. \tag{A.4}$$

Note that, according to the Stochastic Representation Theorem [2], we have that $Q_l^{1/2} =_d \mathcal{R}$ and $Q_l =_d \mathcal{Q}$. Consider now the term $E_p\{\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\}$. By exploiting the expression in (A.4), we have that:

$$E_p\left\{\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\right\} = -E_p\{\phi(G_l^0)Q_l\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\} - E_p\{\phi(G_l^0)Q_l^{1/2}\mathbf{z}_l^H\mathbf{h}_{il}\} - E_p\{\phi(G_l^0)Q_l^{1/2}\mathbf{h}_{il}^H\mathbf{z}_l\} - A_{il}E_p\{\phi(G_l^0)\}. \tag{A.5}$$

It must be noted now that, since G_l^0 has been defined in (20) as,

$$G_l^0 \triangleq (\mathbf{x}_l - \boldsymbol{\gamma}_l^0)^H\boldsymbol{\Pi}_0^{-1}(\mathbf{x}_l - \boldsymbol{\gamma}_l^0) = (Q_l^{1/2}\mathbf{z}_l + \mathbf{r}_l^0)^H\boldsymbol{\Pi}_0^{-1}(Q_l^{1/2}\mathbf{z}_l + \mathbf{r}_l^0) \tag{A.6}$$

the random variable $\phi(G_l^0)$ and the other random quantities $Q_l^{1/2}$ and $\mathbf{z}_l$ are mutually *dependent* and consequently the closed form evaluation of the expectation operator in (A.5) is not feasible. Similar considerations hold true for all the other terms involving the expectation operator w.r.t. the true pdf $p_X(\mathbf{x})$, that are $E_p\left\{\phi(G_l^0)\phi(G_m^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\}$ for $E_p\{V_{ij}(\boldsymbol{\theta}_0)\}$ in (21), $E_p\left\{\phi'(G_l^0)\frac{\partial G_l^0}{\partial\theta_j}\frac{\partial G_l^0}{\partial\theta_i}\right\}$

and $E_p\{\phi(G_l^0)\frac{\partial^2 G_l^0}{\partial\theta_i\partial\theta_j}\}$ for $E_p\{H_{ij}(\boldsymbol{\theta}_0)\}$ in (24). There are, however, two important cases in which $\phi(G_l^0)$ is statistically independent of $Q_l^{1/2}$ (and consequently of Q_l) and $\mathbf{z}_l$: these are the Scenario 1 and the Scenario 2 which are discussed in the following Appendices.

Appendix B. Proof of the SB formulas for the Scenario 1

As discussed in Section 3.2, in the Scenario 1, the true marginal pdf is an unspecified CES pdf $p_X(\mathbf{x}_l) = CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g)$, while the assumed one is a parametric complex Gaussian pdf $f_X(\mathbf{x}_l;\boldsymbol{\theta}) = \mathcal{CN}(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}))$. As a consequence, $\phi(G_l(\boldsymbol{\theta})) \equiv -1, \forall\boldsymbol{\theta}\in\Theta$ and then it trivially satisfy the independence condition discussed in Appendix A. In order to obtain a closed form expression of the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$, the second term in (30) can be rewritten as:

$$\sum_{l=1}^{L}\sum_{m=1}^{L} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\} = \sum_{l=1}^{L} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_l^0}{\partial\theta_j}\right\} + \sum_{l=1}^{L}\sum_{\substack{m=1,\\ m\neq l}}^{L} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\}. \tag{B.1}$$

By using (A.4) and the Stochastic Representation Theorem, when $l\neq m$ we have that:

$$\begin{aligned} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\} &= E_Q^2\{Q\}E_p\left\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\right\}E_p\left\{\mathbf{z}_m^H\mathbf{S}_j^0\mathbf{z}_m\right\} \\ &\quad + E_Q\{Q\}\left(A_{jm}E_p\left\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\right\}\right. \\ &\quad \left. + A_{il}E_p\left\{\mathbf{z}_m^H\mathbf{S}_j^0\mathbf{z}_m\right\}\right) + A_{il}A_{jm}, \end{aligned} \tag{B.2}$$

where A_{il} is defined in (A.3) and where we used the facts that $\mathbf{z}_l$ and $\mathbf{z}_m$ are independent and $E_p\{\mathbf{z}_l\} = E_p\{\mathbf{z}_m\} = \mathbf{0}$. When $l = m$, the term in (B.2) has to be recast as:

$$\begin{aligned} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_l^0}{\partial\theta_j}\right\} &= E_Q\{Q^2\}E_p\left\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\mathbf{z}_l^H\mathbf{S}_j^0\mathbf{z}_l\right\} \\ &\quad + E_Q\{Q\}\left(A_{jl}E_p\left\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\right\} + A_{il}E_p\left\{\mathbf{z}_l^H\mathbf{S}_j^0\mathbf{z}_l\right\}\right. \\ &\quad \left. + E_p\left\{\mathbf{z}_l^H\mathbf{h}_{il}\mathbf{h}_{jl}^H\mathbf{z}_l\right\} + E_p\left\{\mathbf{h}_{il}^H\mathbf{z}_l\mathbf{z}_l\mathbf{h}_{jl}\right\}\right) + A_{il}A_{jl}, \end{aligned} \tag{B.3}$$

where $\mathbf{h}_{il}$ is defined in (A.2) and where we used the fact that the third-order moments of $\mathbf{u}_l$ (and then of $\mathbf{z}_l$) vanish (see Lemma 1 in [2]). Moreover, as a consequence of the circularity property of $\mathbf{z}_l$ [2], the following equalities hold:

$$E_p\{\mathbf{z}_l\mathbf{z}_l^T\} = E_p\{\mathbf{z}_l^*\mathbf{z}_l^H\} = \mathbf{0}, \tag{B.4}$$

$$E_p\{\mathbf{z}_l^H\mathbf{h}_{il}\mathbf{z}_l^H\mathbf{h}_{jl}\} = E_p\{\mathbf{h}_{il}^T\mathbf{z}_l^*\mathbf{z}_l^H\mathbf{h}_{jl}\} = 0, \tag{B.5}$$

$$E_p\{\mathbf{h}_{il}^H\mathbf{z}_l\mathbf{h}_{jl}^H\mathbf{z}_l\} = E_p\{\mathbf{h}_{il}^H\mathbf{z}_l\mathbf{z}_l^T\mathbf{h}_{jl}^*\} = 0. \tag{B.6}$$

From the properties of the trace operator, we have that:

$$E_p\left\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\right\} = N^{-1}\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}), \tag{B.7}$$

$$E_p\left\{\mathbf{z}_l^H\mathbf{h}_{il}\mathbf{h}_{jl}^H\mathbf{z}_l\right\} = N^{-1}\mathrm{tr}(\mathbf{h}_{il}\mathbf{h}_{jl}^H\boldsymbol{\Sigma}), \tag{B.8}$$

$$E_p\left\{\mathbf{h}_{il}^H\mathbf{z}_l\mathbf{z}_l\mathbf{h}_{jl}\right\} = N^{-1}\mathrm{tr}(\mathbf{h}_{jl}\mathbf{h}_{il}^H\boldsymbol{\Sigma}), \tag{B.9}$$

while, by exploiting exactly the same procedure used in [11,20], it can be shown that:

$$E_p\left\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\mathbf{z}_l^H\mathbf{S}_j^0\mathbf{z}_l\right\} = \frac{\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma})\mathrm{tr}(\mathbf{S}_j^0\boldsymbol{\Sigma}) + \mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}\mathbf{S}_j^0\boldsymbol{\Sigma})}{N(N+1)}. \tag{B.10}$$

Another useful relation can be obtained by exploiting the equality in (27). In fact, by combining (A.5) with (27), we have that:

$$E_Q\{Q\}E_p\{\mathbf{z}_l^H\mathbf{S}_i^0\mathbf{z}_l\} = -\mathrm{tr}(\mathbf{P}_i^0) - A_{il}, \quad i = 1,\ldots,d, \tag{B.11}$$

and then, from (B.7) and from the identifiablity condition for CES distribution assumed in this paper, i.e. $E_Q\{Q\} = N$, we obtain:

$$A_{il} = -\mathrm{tr}(\mathbf{P}_i^0) - \mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}), \quad i = 1,\ldots,d. \tag{B.12}$$

Using (B.11) and the condition $E_Q\{Q\} = N$ (and consequently $E_Q^2\{Q\} = N^2$), the term in (B.2) can be easily evaluated as:

$$E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\} = \mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0). \tag{B.13}$$

Similarly, using (B.12), (B.8), (B.9), (B.10) and the condition $E_Q\{Q\} = N$, the term in (B.3) can be expressed as:

$$\begin{aligned} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_l^0}{\partial\theta_j}\right\} &= \mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) + \mathrm{tr}(\mathbf{h}_{il}\mathbf{h}_{jl}^H\boldsymbol{\Sigma}) + \mathrm{tr}(\mathbf{h}_{jl}\mathbf{h}_{il}^H\boldsymbol{\Sigma}) \\ &\quad + \left(\frac{E_Q\{Q^2\}}{N(N+1)} - 1\right)\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma})\mathrm{tr}(\mathbf{S}_j^0\boldsymbol{\Sigma}) \\ &\quad + \mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}\mathbf{S}_j^0\boldsymbol{\Sigma}) \end{aligned} \tag{B.14}$$

Finally, by combining all the previous results, we have that:

$$\begin{aligned} \sum_{l=1}^{L}\sum_{m=1}^{L} E_p\left\{\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\} &= L^2\mathrm{tr}(\mathbf{P}_i^0)\mathrm{tr}(\mathbf{P}_j^0) \\ &\quad + \sum_{l=1}^{L}\left[\mathrm{tr}(\mathbf{h}_{il}\mathbf{h}_{jl}^H\boldsymbol{\Sigma}) + \mathrm{tr}(\mathbf{h}_{jl}\mathbf{h}_{il}^H\boldsymbol{\Sigma})\right] \\ &\quad + L\left(\frac{E_Q\{Q^2\}}{N(N+1)} - 1\right)\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma})\mathrm{tr}(\mathbf{S}_j^0\boldsymbol{\Sigma}) \\ &\quad + L\frac{E_Q\{Q^2\}}{N(N+1)}\mathrm{tr}(\mathbf{S}_i^0\boldsymbol{\Sigma}\mathbf{S}_j^0\boldsymbol{\Sigma}), \end{aligned} \tag{B.15}$$

from which (31) follows immediately. Regarding the calculation of the matrix $\mathbf{A}(\boldsymbol{\theta}_0)$ in (34), no particular simplification can be made. Specifically, the term $E_p\left\{\frac{\partial^2 G_l^0}{\partial\theta_i\partial\theta_j}\right\}$ in (33) has to be obtained through direct calculation from (A.1). Since the derivation is long, tedious and does not add any insightful considerations about the problem at hand, we decided to not report it here.

Appendix C. Proof of the SB formulas for the Scenario 2

As previously discussed in the paper, in the Scenario 2 we suppose that the true marginal pdf is given by a parametric CES distribution such that $p_X(\mathbf{x}_l) \triangleq p_X(\mathbf{x}_l;\bar{\boldsymbol{\theta}}) = CES_N(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}, g)$, where $\boldsymbol{\mu}_l = \boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}})$ and $\boldsymbol{\Sigma} = \boldsymbol{\Pi}(\bar{\boldsymbol{\theta}})$ for a given $\bar{\boldsymbol{\theta}}\in\Theta$. The assumed pdf is itself a parametric CES distribution that share the same parametrization of the true one but may have a different density generator, i.e. $f_X(\mathbf{x}_l;\boldsymbol{\theta}) = CES_N(\boldsymbol{\gamma}_l(\boldsymbol{\theta}), \boldsymbol{\Pi}(\boldsymbol{\theta}), w)$, $\boldsymbol{\theta}\in\Theta$, and possibly $g(t)\neq w(t), \forall t\in\mathbb{R}^+$. In order to guarantee the correct identifiability of the true and the assumed CES distributions, as before, we may impose a constraint of both the modular variate, that is:

$$\begin{aligned} &E_p\{(\mathbf{x}_l - \boldsymbol{\mu}_l)^H\boldsymbol{\Sigma}^{-1}(\mathbf{x}_l - \boldsymbol{\mu}_l)\} \triangleq E_p\{(\mathbf{x}_l - \boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}}))^H\boldsymbol{\Pi}(\bar{\boldsymbol{\theta}})^{-1} \\ &\times(\mathbf{x}_l - \boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}}))\} = E_Q\{Q_l\} = N, \end{aligned} \tag{C.1}$$

$$E_{f_{\boldsymbol{\theta}}}\{(\mathbf{x}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}))^H\boldsymbol{\Pi}(\boldsymbol{\theta})^{-1}(\mathbf{x}_l - \boldsymbol{\gamma}_l(\boldsymbol{\theta}))\} = N, \forall\boldsymbol{\theta}\in\Theta. \tag{C.2}$$

Let us suppose now that, for a given true pdf $p_X(\mathbf{x}_l) \triangleq p_X(\mathbf{x}_l;\bar{\boldsymbol{\theta}})$ and for a given assumed pdf $f_X(\mathbf{x}_l;\boldsymbol{\theta})$, the pseudo-true parameter vector $\boldsymbol{\theta}_0$ equates the true parameter vector $\bar{\boldsymbol{\theta}}$, and consequently, from (14):

$$\left.\frac{\partial D(p_X\|f_{\boldsymbol{\theta}})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}} = -E_p\left\{\left.\frac{\partial \ln f_X(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i}\right|_{\boldsymbol{\theta}=\bar{\boldsymbol{\theta}}}\right\} = 0, \quad i = 1,\ldots,d. \tag{C.3}$$

Under this assumption, we have that:

$$\begin{aligned} G_l^0 &\triangleq (\mathbf{x}_l - \boldsymbol{\gamma}_l^0)^H \boldsymbol{\Pi}_0^{-1} (\mathbf{x}_l - \boldsymbol{\gamma}_l^0) \\ &= (\mathbf{x}_l - \boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}}))^H \boldsymbol{\Pi}(\bar{\boldsymbol{\theta}})^{-1} (\mathbf{x}_l - \boldsymbol{\gamma}_l(\bar{\boldsymbol{\theta}})) \\ &= (\mathbf{x}_l - \boldsymbol{\mu}_l)^H \boldsymbol{\Sigma}^{-1} (\mathbf{x}_l - \boldsymbol{\mu}_l) \triangleq Q_l =_d \mathcal{Q} = \mathcal{R}^2, \end{aligned} \tag{C.4}$$

and consequently, the Stochastic Representation Theorem allows us to write the following equality:

$$\mathbf{x}_l - \boldsymbol{\gamma}_l^0 = \mathbf{x}_l - \boldsymbol{\mu}_l =_d \mathcal{R}\mathbf{T}\mathbf{u}_l =_d \mathcal{R}\boldsymbol{\Pi}_0^{1/2}\mathbf{u}_l. \tag{C.5}$$

where $\boldsymbol{\gamma}_l^0 \triangleq \boldsymbol{\gamma}_l(\boldsymbol{\theta}_0)$ and $\boldsymbol{\Pi}_0 \triangleq \boldsymbol{\Pi}(\boldsymbol{\theta}_0)$. It is worth to highlight that the equality chains in (C.4) and (C.5) hold true if and only if $\boldsymbol{\theta}_0 = \bar{\theta}$. We are quite confident that, in the context of Scenario 2 (i.e. when the misspecification is only due to a wrong assumption on the density generator), the equality $\boldsymbol{\theta}_0 = \bar{\theta}$ always holds true. Since we have not a proof of this fact yet, we considered it as an assumption (Assumption A2 in Section 3.1).

As a consequence of (C.4), the expectation of functions of G_l^0, say $h(G_l^0)$, can be explicitly derived as:

$$\begin{aligned} E_{\mathcal{Q}}\{h(G_l^0)\} &\triangleq \int_0^\infty h(t)\, p_{\mathcal{Q}}(t)dt = \delta_{N,g}^{-1} \int_0^\infty h(t) t^{N-1} g(t)dt \\ &= E_{\mathcal{Q}}\{h(Q)\}, \forall l \in \mathbb{N}. \end{aligned} \tag{C.6}$$

Since such expectation does not depend on the index l and on the unknown parameter vector $\boldsymbol{\theta}$, in the rest of this Appendix we always indicate $E_{\mathcal{Q}}\{h(G_l^0)\}$ simply as $E_{\mathcal{Q}}\{h(Q)\}$.

As in Appendix B for the Scenario 1, in order to obtain a closed form expression of the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$, the second term in (28) can be rewritten as:

$$\begin{aligned} &\sum_{l=1}^{L}\sum_{m=1}^{L} E_p\left\{\phi(G_l^0)\phi(G_m^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\} \\ &= \sum_{l=1}^{L} E_p\left\{\phi^2(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_l^0}{\partial\theta_j}\right\} \\ &+ \sum_{l=1}^{L}\sum_{\substack{m=1,\\ m\neq l}}^{L} E_p\left\{\phi(G_l^0)\phi(G_m^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\}. \end{aligned} \tag{C.7}$$

The term $\frac{\partial G_l^0}{\partial\theta_i}$ can be easily obtained by firstly taking the derivative of the terms in (22) and then by substituting in the obtained expression the Stochastic Representation of the difference vector $\mathbf{x}_l - \boldsymbol{\gamma}_l^0$ in (C.5), we have that:

$$\frac{\partial G_l^0}{\partial\theta_i} = -Q_l^{1/2}\left(\left(\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H \boldsymbol{\Pi}_0^{-1/2}\mathbf{u}_l + \mathbf{u}_l^H\boldsymbol{\Pi}_0^{-1/2}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right) - Q_l\mathbf{u}_l^H\mathbf{P}_i^0\mathbf{u}_l. \tag{C.8}$$

Let us now recall from [18] the following three equalities:

$$E_p\{\mathbf{u}_l^H\mathbf{P}_i\mathbf{u}_l\mathbf{u}_l^H\mathbf{P}_j\mathbf{u}_l\} = \frac{\mathrm{tr}(\mathbf{P}_i)\mathrm{tr}(\mathbf{P}_j) + \mathrm{tr}(\mathbf{P}_i\mathbf{P}_j)}{N(N+1)} \tag{C.9}$$

and

$$E_p\{\mathbf{u}_l^H\mathbf{P}_i\mathbf{u}_l\} = E_p\{\mathrm{tr}(\mathbf{P}_i\mathbf{u}_l\mathbf{u}_l^H)\} = N^{-1}\mathrm{tr}(\mathbf{P}_i) \tag{C.10}$$

and

$$E_p\{(\mathbf{u}_l^H\mathbf{a})(\mathbf{u}_l^H\mathbf{D}\mathbf{u}_l)\} = 0 \tag{C.11}$$

for some vector $\mathbf{a}$ and hermitian matrix $\mathbf{D}$.

By substituting (C.8) in (C.7) and by exploiting the equalities (C.9)–(C.11), and remembering that $E_p\{\mathbf{u}_l\mathbf{u}_l^H\} = (1/N)\mathbf{I}$ (see Section 2), we get:

$$\begin{aligned} &\sum_{l=1}^{L}\sum_{m=1}^{L} E_p\left\{\phi(G_l^0)\phi(G_m^0)\frac{\partial G_l^0}{\partial\theta_i}\frac{\partial G_m^0}{\partial\theta_j}\right\} \\ &= \frac{2E_{\mathcal{Q}}\{Q\phi^2(Q)\}}{N}\sum_{l=1}^{L}\mathrm{Re}\left[\left(\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_j}\right] \\ &+ L\frac{E_{\mathcal{Q}}\{Q^2\phi^2(Q)\}}{N(N+1)}\left(\mathrm{tr}(\mathbf{P}_i)\mathrm{tr}(\mathbf{P}_j) + \mathrm{tr}(\mathbf{P}_i\mathbf{P}_j)\right) \\ &+ L(L-1)\frac{E_{\mathcal{Q}}^2\{Q\phi(Q)\}}{N^2}\mathrm{tr}(\mathbf{P}_i)\mathrm{tr}(\mathbf{P}_j). \end{aligned} \tag{C.12}$$

Let us now impose the equality in (27). In particular, we have that:

$$\begin{aligned} &\sum_{l=1}^{L} E_p\left\{\phi(G_l^0)\frac{\partial G_l^0}{\partial\theta_i}\right\} = -E_{\mathcal{Q}}\{Q\phi(Q)\}\sum_{l=1}^{L}E_p \\ &\times\left\{\left(\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right)^H\boldsymbol{\Pi}_0^{-1/2}\mathbf{u}_l + \mathbf{u}_l^H\boldsymbol{\Pi}_0^{-1/2}\frac{\partial\boldsymbol{\gamma}_l^0}{\partial\theta_i}\right\} \\ &-E_{\mathcal{Q}}\{Q\phi(Q)\}\sum_{l=1}^{L}E_p\left\{\mathbf{u}_l^H\mathbf{P}_i^0\mathbf{u}_l\right\} \\ &= -LN^{-1}E_{\mathcal{Q}}\{Q\phi(Q)\}\mathrm{tr}(\mathbf{P}_i), \end{aligned} \tag{C.13}$$

where we used the fact that the fact that $E_p\{\mathbf{u}_l\} = \mathbf{0}$ while the last equality follows from (C.10). Then, by exploiting the equality (27), we obtain the following relation:

$$E_{\mathcal{Q}}\{Q\phi(Q)\} = -N, \tag{C.14}$$

where $\phi(t) = w'(t)/w(t)$ and $w(t)$ is the density generator of the assumed CES distribution $f_X(\mathbf{x}_l; \boldsymbol{\theta})$. Finally, by substituting (C.14) in (C.12), and then replacing the obtained term in (28), we get the closed form expression of the matrix $\mathbf{B}(\boldsymbol{\theta}_0)$ given in (35). The evaluation of the matrix $\mathbf{A}(\boldsymbol{\theta}_0)$ in (36) follows directly from the direct calculation of the terms $E_p\left\{\phi'(G_l^0)\frac{\partial G_l^0}{\partial\theta_j}\frac{\partial G_l^0}{\partial\theta_i}\right\}$ and $E_p\left\{\phi(G_l^0)\frac{\partial^2 G_l^0}{\partial\theta_i\partial\theta_j}\right\}$. In particular, all the derivatives have to be evaluated from (22), while, to evaluate the expectation operator w.r.t. the true distribution $p_X(\mathbf{x})$, one has to use the Stochastic Representation in (C.5) by keeping in mind that $\bar{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$. Since this calculation is tedious and not informative, we decided to not report it here.

References

[1] A.M. Zoubir, V. Koivunen, Y. Chakhchoukh, M. Muma, Robust estimation in signal processing: a tutorial-style treatment of fundamental concepts, IEEE Signal Process. Mag. 29 (4) (2012) 61–80.

[2] E. Ollila, D.E. Tyler, V. Koivunen, H.V. Poor, Complex elliptically symmetric distributions: survey, new results and applications, IEEE Trans. Signal Process. 60 (11) (2012) 5597–5625.

[3] C.D. Richmond, Adaptive Array Signal Processing and Performance Analysis in Non-Gaussian Environments, Massachusetts Institute of Technology, 1996, Ph.D. thesis. Available at: https://dspace.mit.edu/handle/1721.1/11005

[4] P.J. Huber, The behavior of maximum likelihood estimates under nonstandard conditions, in: Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability, Univ. of California, Berkeley, CA, USA, 1967.

[5] H. White, Maximum likelihood estimation of misspecified models, Econometrica (1982) 1–25.

[6] Q.H. Vuong, Cramér-rao Bounds for Misspecified Models, Div. of the Humanities and Social Sci., California Inst. of Technol., Pasadena, CA, USA, 1986. Working Paper 652, Oct. [Online]. Available: https://www.hss.caltech.edu/content/cramer-rao-bounds-misspeci-fied-models

[7] C.D. Richmond, L.L. Horowitz, Parameter bounds on estimation accuracy under model misspecification, IEEE Trans. Signal Process. 63 (9) (2015) 2263–2278.

[8] S. Fortunati, F. Gini, M. Greco, The Misspecified Cramér-Rao bound and its application to scatter matrix estimation in complex elliptically symmetric distributions, IEEE Trans. Signal Process. 64 (9) (2016) 2387–2399.

[9] S. Fortunati, F. Gini, M.S. Greco, The constrained Misspecified Cramér-Rao bound, IEEE Signal Process. Lett. 23 (5) (2016) 718–721.

[10] C. Ren, M.N. El Korso, J. Galy, E. Chaumette, P. Larzabal, A. Renaux, Performance bounds under misspecification model for mimo radar application, in: 2015 23rd European Signal Processing Conference (EUSIPCO), 2015, pp. 514–518.

[11] M. Greco, S. Fortunati, F. Gini, Maximum likelihood covariance matrix estimation for complex elliptically symmetric distributions under mismatched conditions, Signal Process. 104 (2014) 381–386.

[12] S. Fortunati, F. Gini, M.S. Greco, Matched, mismatched, and robust scatter matrix estimation and hypothesis testing in complex t-distributed data, EURASIP J. Adv. Signal Process. 2016 (1) (2016) 123.
[13] C.D. Richmond, P. Basu, Bayesian framework and radar: on misspecified bounds and radar-communication cooperation, in: 2016 IEEE Statistical Signal Processing Workshop (SSP), 2016, pp. 1–4.
[14] P.A. Parker, C.D. Richmond, Methods and bounds for waveform parameter estimation with a misspecified model, in: 2015 49th Asilomar Conference on Signals, Systems and Computers, 2015, pp. 1702–1706.
[15] S.M. Kay, Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory, 1993.
[16] D. Slepian, Estimation of signal parameters in the presence of noise, Trans. IRE Prof. Group Inf. Theory 3 (3) (1954) 68–89.
[17] W.J. Bangs, Array Processing with Generalized Beamformers, Ph.D. thesis, Yale University, New Haven, CT, USA, 1971.
[18] O. Besson, Y.I. Abramovich, On the fisher information matrix for multivariate elliptically contoured distributions, IEEE Signal Process. Lett. 20 (11) (2013) 1130–1133.
[19] S. Fortunati, Misspecified Cramér-Rao bounds for complex unconstrained and constrained parameters, in: 2017 25th European Signal Processing Conference (EUSIPCO), 2017.
[20] M. Greco, F. Gini, Cramér-Rao lower bounds on covariance matrix estimation for complex elliptically symmetric distributions, IEEE Trans. Signal Process. 61 (24) (2013) 6401–6409.